Derivatives of the fixed-node energy

S. De Palo[(a,b)], S. Moroni[(a,b)] and S. Baroni[(a,c)]
[a] *Istituto Nazionale per la Fisica della Materia*
[b] *Dipartimento di Fisica,*
Università di Roma "La Sapienza"
[c] *Scuola Internazionale Superiore di Studi Avanzati, Trieste*
Italy

We present quantum Monte Carlo calculations of total energy derivatives, consistently performed in the fixed-node approximation. Contributions from nodal displacements, neglected or approximated in previous investigations, are properly taken into account. Their impact on the efficiency is discussed.

I. INTRODUCTION

Quantum Monte Carlo simulations are emerging as a promising alternative route towards high-accuracy predictions of molecular and materials properties. The fixed-node approximation[1] is widely used in quantum Monte Carlo simulations[2] of many-fermion systems to avoid the sign problem[3]. It consists of solving for the lowest-energy eigenstate, Φ, of the Hamiltonian, H, with the boundary conditions that the wave function vanishes on the nodal surface of an anti-symmetric trial function, Ψ, i.e. at the points R in configuration space where $\Psi(R) = 0$. The fixed-node energy (FNE)—*i.e.* the eigenvalue E of the Hamiltonian corresponding to Φ—is an upper bound for the ground state energy, and it becomes exact if Ψ has the same nodes as the true ground state wave function, Φ_{ex}. Therefore, for a given trial function, the FNE is more accurate than the variational energy (the expectation value of H on Ψ), because exactness of the latter requires the stronger condition that Ψ coincide with Φ_{ex} for all R. Indeed, the fixed-node approximation offers the best viable numerical solution for several many-fermion problems[2].

The ability to calculate low-order derivatives of the Born-Oppenheimer, energy with respect to an external parameter, λ, is an important ingredient of the calculation of materials properties. Forces acting on individual atoms or the stress acting within an extended system are examples of first-order derivatives whose vanishing determines the mechanical equilibrium of a system. Similarly, vibrational frequencies—as well as any kind of static generalized susceptibility—are determined by second-order derivatives. The calculation of derivatives of the FNE, $\partial E/\partial\lambda$, has the potential of largely extending the scope of applications of quantum Monte Carlo simulations and spawning further methodological developments (besides the examples given above, λ could also be the strength of any external field, or a variational parameter which modifies the nodal surface of Ψ). However technical difficulties[4] have long hindered routine calculation of FNE derivatives, and despite recent progress[5,6,7] open questions still remain. In particular, the nodal displacement upon variation of λ has always been neglected *in toto* or in part[8], thus adding a further bias on top of the fixed-node approximation: referring to recent literature, λ-independent nodes have been chosen in the path integral Monte Carlo calculation of electronic forces[5]; they *have* to be chosen so in applications of the Hellmann-Feynman theorem[7,9]; finally, the drift-diffusion factor of the propagator, together with its nodal contributions, has been dropped from the weights in correlated-sampling quantum Monte Carlo simulations[6].

In this paper we address the calculation of the derivatives of the FNE without any further approximations. To this end, we examine in detail a simple but significant model that, we believe, is representative of the difficulties which are met in the calculation of these derivatives in more realistic situations. The focus being on the role of nodal displacements, we consider a system where such a role is crucial and a check against exact results is possible, namely a spinless Fermi gas in two dimensions in a weak external field.

II. METHODOLOGICAL ASPECTS AND THE MODEL

The Hamiltonian of our model system is:

$$H = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 + \lambda \sum_{i=1}^{N} \cos(\mathbf{q}\cdot\mathbf{r}_i), \qquad (1)$$

where the strength of the external field, λ, is the parameter upon which the FNE is considered to depend, and where periodic boundary conditions have been assumed. The ground state is a Slater determinant, D, of λ-dependent Mathieu functions[13], whereas our trial function is $\Psi = J \times D$ with a symmetric Jastrow factor $J(R) = \exp(-\sum_{i,j} u(r_{ij}))$ which introduces spurious pair correlations. The fixed-node solution, exact in this case, is recovered as an absolutely non-trivial result of the simulation, since the trial function, wrong on purpose, can be made poor at will by tuning the *pseudo-potential*, u.

We use the reptation quantum Monte Carlo (RQMC) method[14], in which the calculation of derivatives is simpler to implement than with branching algorithms—just as simple as in path integral Monte Carlo[5]. The FNE

can be computed from the mixed estimate of the Hamiltonian,

$$E = \frac{\langle\Psi|H|\Phi\rangle}{\langle\Psi|\Phi\rangle} = \frac{\int E_L(R) f(R) dR}{\int f(R) dR}, \qquad (2)$$

by averaging the local energy $E_L(R) = H\Psi(R)/\Psi(R)$ over the mixed distribution $f(R) = \Phi(R)\Psi(R)$. The unknown state Φ is expressed in terms of Ψ through the imaginary-time propagator, $\Phi = \lim_{\tau\to\infty} e^{-H\tau}\Psi$. Upon discretization of the imaginary time evolution and introduction of importance sampling, the mixed distribution can be cast into the form:

$$f(R_M) = \lim_{M\to\infty} \int \Pi_{i=0}^{M-1} G(R_i, R_{i+1}; \epsilon)\Psi(R_0)^2, \qquad (3)$$

where $\epsilon = \tau/M$ is the time step. The importance-sampled Green's function, $G(R, R'; \epsilon) = \Psi(R)\langle R|e^{-\epsilon H}|R'\rangle/\Psi(R')$, is replaced by a short-time approximation, a common choice being[2]

$$G_F(R, R'; \epsilon) \propto e^{-(R'-R-\epsilon F(R)/2)^2/2\epsilon} e^{-\frac{\epsilon}{2}(E_L(R)+E_L(R'))}, \qquad (4)$$

with $F = \nabla \ln \Psi^2$.

In the RQMC method a generalized Metropolis algorithm is used to sample the probability distribution for a path of finite length in imaginary time (a *reptile*, $X = \{R_0, \ldots, R_M\}$,

$$P(X) = \Pi_{i=0}^{M-1} G(R_i, R_{i+1}; \epsilon)\Psi(R_0)^2, \qquad (5)$$

the limit in Eq. (3) being reached for large enough M. Replacement of Eq. (3) in (2) shows that the FNE can be evaluated by averaging the local energy on the distribution P,

$$E = \langle\langle E_L(R_M)\rangle\rangle \equiv \frac{\int dX E_L(R_M) P(X)}{\int dX P(X)}. \qquad (6)$$

Since $P(X)$ is explicitly known, we can differentiate to get

$$\begin{aligned} \partial_\lambda E &= \langle\langle A(X)\rangle\rangle, \qquad (7)\\ A(X) &= [\partial_\lambda E_L(R_M) + (E_L(R_M) - E)\partial_\lambda \ln P(X)] \end{aligned}$$

Finite increments of amplitude Δ are used for the evaluation of $\partial_\lambda E_L(R)$ and $\partial_\lambda \ln P(X)$ on the path configurations sampled from $P(X)$. Equivalently, finite increments can be used for E via the correlated-sampling method:

$$\begin{aligned} E' &= \frac{\int dX P'(X) E'_L(R_M)}{\int dX P'(X)} \qquad (8)\\ &= \frac{\int dX P(X) W(X) E'_L(R_M)}{\int dX P(X) W(X)}, \end{aligned}$$

with $W(X) = P'(X)/P(X)$, primed quantities being evaluated at $\lambda + \Delta$.

In the *variational* Quantum Monte Carlo (VMC) approximation, the probability distribution $P(X)$ in Eq. (8) or (9) is replaced by $\Psi^2(R_M)$. Small energy differences are routinely calculated by VMC[4]. The more accurate FNA requires instead the full probability distribution of Eq (5), which entails nontrivial difficulties through the term $\partial_\lambda \ln P(X)$ of Eq. (8). Therefore we now examine in some detail the behavior of the probability distribution P when a configuration along the path, say R_i, approaches a node of Ψ.

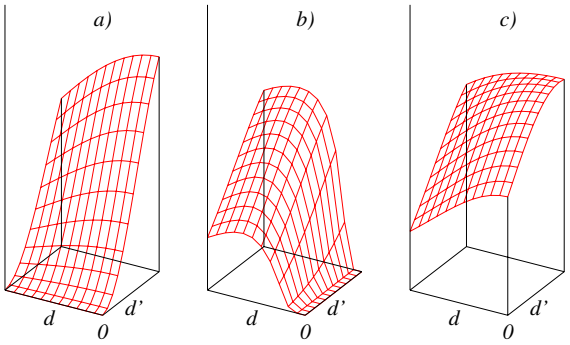

FIG. 1: Importance-sampled Green's function for a one-dimensional motion near a hard wall. Left panel: exact; middle panel: Eq. (4); right panel: Eq. (4) with a smoothed F. The range of d and d' is $(0, \epsilon^{1/2})$.

When the distance d from the nodal surface is small enough, the dynamics of the system is dominated by the hard wall potential due to the fixed-node boundary conditions[11]. The importance-sampled Green's function should behave like that of a one-dimensional motion close to a hard wall $[G_0(d, d'; \epsilon) \propto \frac{d'}{d} e^{-(d-d')^2/2\epsilon}(1 - e^{-2dd'/\epsilon})]$, as shown in the left panel of Fig. (1). The short time approximation $G_F(R, R', \epsilon)$, instead, is a very poor approximation near a node, because it assumes $F = \nabla \ln \Psi^2$ to be constant during a time step, whereas F diverges as d^{-1}. The approximation (4) for a one-dimensional motion close to a hard wall, shown in the middle panel of Fig. (1), also bears little resemblance to G_0. First, G_F does not vanish at $d = 0$ where G_0 does. Sampling from G_F would thus imply an unphysical finite density at the nodes; the correct density, which is proportional to d^2 can be recovered including in the algorithm a Metropolis test based on the square of the trial function[1,12]. Second, G_F vanishes at $d' = 0$ where G_0 does not. This would imply a divergence proportional to d^{-2} of $\ln P$ at the nodes, thus making the evaluation of the derivative, Eq. (8), impossible. Such a divergence can be eliminated, leaving the zero-time-step limit unchanged, by a suitable cutoff on the quantum force, $\nabla \ln \Psi$. We thus replace F in Eq. 4 with $\bar{F} = 4F(-1 + \sqrt{1 + F^2\epsilon/2})/(F^2\epsilon)$[12]. The resulting Green's function, $G_{\bar{F}}$, is shown in the right panel of Fig. (1). With this replacement the variance of the estimator (8) of the derivative is finite. However we

empirically find that the extrapolation of the result at $\epsilon \to 0$ can be problematic (see also Ref. [10])

In the simulation of electrons in atoms and molecules, it proves convenient to use an interpolation between Eq. (4) and a locally better Green's function near the nuclei[12]. In the same spirit, we use an interpolation $\tilde{G}$ between $G_{\bar{F}}$ and a locally better Green's function G_N near the nodes:

$$\tilde{G}(R,R';\epsilon) = \alpha(d)G_{\bar{F}}(R,R';\epsilon) + (1-\alpha(d))G_N(R,R';\epsilon) \quad (9)$$

where the function $\alpha(d)$ smoothly varies from 0 to 1 as d goes from 0 to $\sim \epsilon^{1/2}$. In the new Green's function

$$\begin{aligned} G_N(R,R';\epsilon) \;&\propto\; \frac{\Psi(R')}{\Psi(R)} e^{(R'-R)^2/4\epsilon}(1-e^{-d'd/\epsilon}) \quad (10)\\ &\times\; e^{-\epsilon(V(R')+V(R))/2} \end{aligned}$$

the potential V is treated in the primitive approximation, whereas the nodal dependence has the correct behavior of G_0. Although the variance due to $\partial_\lambda \ln P(X)$ is finite using $\tilde{G}$, a time-step dependent cutoff on d is legitimate and can be useful.

III. RESULTS

We turn now to the discussion of our results for the Fermi gas. We first choose a very poor trial function Ψ_I with a pseudo-potential, $u = a\exp(-br^2)$ where $a = 0.9$ and $b = 0.7$, yielding a variational energy about 20% higher than the exact ground state energy (whereas the FNE and the exact result coincide in this case, since the Slater determinant is the true ground state wave function). The derivative $\partial_\lambda E|_{x_0}$, with $\lambda_0 = 0.02$ and $q = 1.585$, is calculated using $\tilde{G}$. The result, shown in Tab. I, does indeed recover the correct value.

For comparison, we also perform calculations using the approximation of Ref. [6], which drops the drift-diffusion part of the propagator in Eq. (8) and introduces an extra factor $\Psi^2(R_M)$ in order to recover the exact result in the limit of perfect importance sampling. As shown in Tab. I, the approximate algorithm is more efficient (it gives a statistical error 3 times smaller for the same number of path configurations). However its result for the derivative is clearly biased. Using a better trial function Ψ_{II} with $a = 0.45$ the systematic bias of the approximate algorithm becomes smaller than statistical error. It is interesting to note that the ratio of the statistical error between the two methods does not seem to depend strongly on the quality of Ψ. We also have evidence that the efficiency depends on the number of particles only through the extra cost of sampling configurations for a larger system (simulations have been performed for $5, 9$ and 21 particles).

In Fig. 2 we show the time-step dependence of various derivatives calculated using different Green's functions. Data obtained with $G_{\bar{F}}$ have smaller statistical errors than with $\tilde{G}$, but their extrapolation to zero time-step is harder to guess in advance. The upper panel shows derivatives of the FNE, $\partial_\lambda E|_{\lambda=0.002}$. The calculation with $\tilde{G}$ extrapolates to the exact result assuming a quadratic time-step dependence, whereas an $\epsilon^{1/2}$ term seems necessary in order to account for the $G_{\bar{F}}$ data.

TABLE I: First derivative of fixed-node energy calculated with Variational Monte Carlo (VMC), using RQMC with $\tilde{G}$ and using the approximated algorithm as explained in the text

Exact	VMC	RQMC	App. Alg.	wave-function
1	0.910(16)	0.998(38)	0.941(14)	Ψ_I
1	0.9677(92)	1.038(38)	0.997(13)	Ψ_{II}

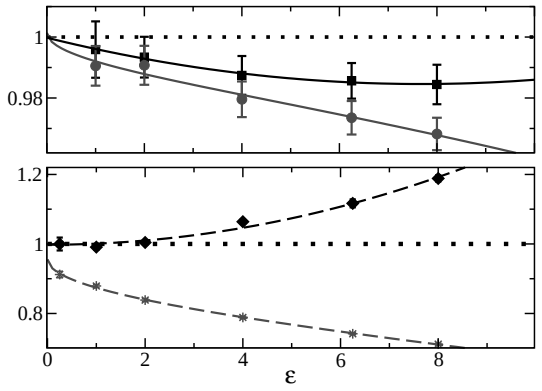

FIG. 2: Time step dependence of the first derivative of the FNE ($\partial_\lambda E|_{x=0.002}$), upper panel, and of the density fluctuation ($\partial_\lambda \rho_{\mathbf{q}}|_{\lambda=0}$), lower panel. We show calculations using $\tilde{G}$ (black squares and diamonds) and $G_{\bar{F}}$ (gray circles and stars); the black lines are quadratic fits, whereas the gray lines include a square root term.

The problems with the $\epsilon \to 0$ limit using $G_{\bar{F}}$ are more evident for the derivative of the density fluctuation, $\partial_\lambda \rho_{\mathbf{q}}|_{\lambda=0}$, where $\rho_{\mathbf{q}}(R) = \sum_i \exp(-i\mathbf{q}\cdot\mathbf{r}_i)$: as we can see from the lower panel of Fig. (2), in this case an even smaller power of ϵ would be required to extrapolate to the exact result at $\epsilon = 0$.

The physical interest in $\partial_\lambda \rho_{\mathbf{q}}|_{\lambda=0}$ stems from its relation to the static dielectric response. The derivative of $\rho_{\mathbf{q}}$ is evaluated as $\langle\langle B(X)\rangle\rangle$, with $B(X) = \partial\rho_{\mathbf{q}}(R_{M/2}) + \rho_{\mathbf{q}}(R_{M/2})\partial_\lambda \ln P(X)$ (cfr. Eq. (8)[15]). The same physical information can be formally obtained from $\partial^2_\lambda E|_{\lambda=0}$. Second derivatives are in principle straightforward to evaluate, but differentiation of Eq. (8) leads to terms containing $\partial^2_\lambda \ln P$ and $(\partial_\lambda \ln P)^2$ which have large fluctuations.

Despite its low efficiency, the calculation of second derivatives of the energy offers an interesting conceptual advantage over the first derivative of the density fluctuation. Suppose we don't know the exact nodal dependence of Ψ on the strength λ of the external field. Starting from

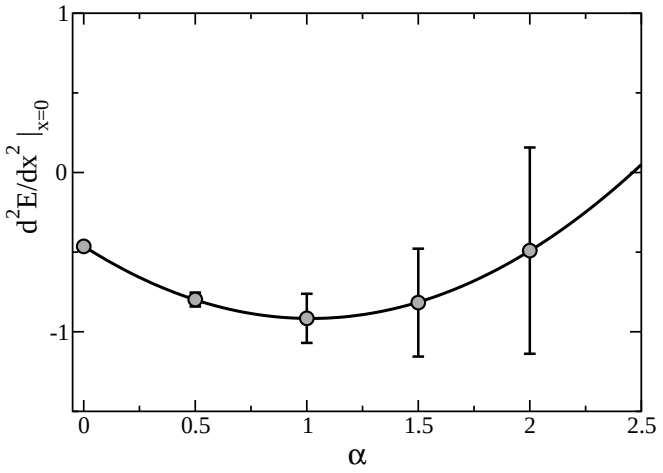

FIG. 3: Second derivative of the FNE, $\partial^2_\lambda E|^{(\alpha)}_{\lambda=0}$ as a function of α. The parameter of the pseudo-potential is $a = 0.90$, and the calculation uses the Green's function $\tilde{G}$. The exact result corresponds to $\alpha = 1$, $\partial^2_\lambda E|^{(\alpha=1)}_{\lambda=0} = -1$.

plane-wave nodes at $\lambda = 0$, we can explore a family of wave functions, Ψ_α, in which the one-particle orbitals are Mathieu functions calculated for an external potential of strength $\alpha\lambda$ *different* from the physical value λ. The nodal dependence of Ψ_α on λ varies with α: for $\alpha = 0$, no nodal displacements are allowed; for $\alpha = 1$, the exact nodal dependence is recovered. We can compute within the *same* run (hence in a correlated way) the values of $\partial^2_\lambda E|^{(\alpha)}_{\lambda=0}$ pertaining to several choices of α. The result is shown in Fig. (3). Because of the variational character of the fixed node approximation, the optimal nodal dependence (within the family Ψ_α, and within error bars) is given by the value α corresponding to the minimum of $\partial^2_\lambda E|^{(\alpha)}_{\lambda=0}$. Inspection of Fig. (3) leads us to conclude that: (i) suppression of the nodal dependence ($\alpha = 0$) leads to a small statistical error, but produces a large bias; (ii) the minimum of $\partial^2_\lambda E|^{(\alpha)}_{\lambda=0}$ is consistent with the exact nodal dependence ($\alpha = 1$); (iii) the dependence of $\partial^2_\lambda E|^{(\alpha)}_{\lambda=0}$ on α is quadratic, so that calculations for just 3 values of α would suffice (this could be useful for a higher-dimensional optimization).

Unfortunately, the calculation of second derivatives has still very large statistical errors (if nodal displacement is properly included). We hope that a successful implementation of noise reduction techniques[16] could eventually turn it into a convenient computational tool. Further perspectives in nodal optimization are offered by the calculation of first derivatives of the FNE, for instance in conjunction with the stochastic gradient approximation method[17].

[1] J.B. Anderson, J. Chem. Phys. **63**, 1499 (1975); *ibid.* **65**, 4121 (1976); J.W. Moskowitz, K. E. Schmidt, M.A. Lee and M.H. Kalos J. Chem. Phys. **77** 349 (1982): P.J. Reynolds , D. M. Ceperley, B. J. alder and W.A. Lester, J. Chem. Phys. **77**, 5593 (1982).

[2] For a recent review on fixed-node diffusion quantum Monte Carlo and further references, see M. Foulkes, L. Mitas, R. Needs and G. Rajagopal, Rev. Mod. Phys. **73**, 33-83 (2001).

[3] M.H. Kalos and P. Whitlock *Monte Carlo Methods* (Wiley & Sons, 1986)

[4] B. L. Hammond, W. A. Lester and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

[5] F. Zong and D. M. Ceperley, Phys. Rev. E **58**, 5123 (1998).

[6] C. Filippi and C. J. Umrigar, Phys. Rev. B **61**, R16291 (2000).

[7] R. Assaraf and M. Caffarel, J. Chem. Phys. **113**, 4028 (2000).

[8] This is true with one exception[10]; however it seems fair to say that the results reported in Ref.[10] need a somewhat educated guess for the extrapolation to zero time step.

[9] F. Schautz and H.-J. Flad, J. Chem. Phys. **110**, 11700 (1999); K. C. Huang, R. J. Needs and G. Rajagopal, J. Chem. Phys. **112**, 4419 (2000); F. Schautz and H.-J. Flad, J. Chem. Phys. **112**, 4421 (2000). Phys. Rev. Lett. **83**, 4682 (1999).

[10] J. Vrbik and S. M. Rothstein, J. Chem. Phys. **96**, 2071 (1992).

[11] We assume in the argument smoothness of the factor $e^{-\frac{\epsilon}{2}(E_L(R)+E_L(R'))}$ of G. It is common practice to smooth the divergence of the local energy at the nodes when evaluating Eq. (4).

[12] C. J. Umrigar, M. P. Nightingale and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).

[13] I. S. Gradshstein and I.M. Ryzhik, *Tables of integrals, series and products* (Academic Press, New York, 1980)

[14] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999)

[15] Within RQMC it is straightforward to compute unbiased expectation values of observables that do not commute with the hamiltonian[14].

[16] See e.g. R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999)

[17] A. Harju, B. Barbiellini, S. Siljamaki, R. M. Nieminen, and G. Ortiz, Phys. Rev. Lett. **79**, 1173-1177 (1997).